DOCUMENT RESUME

ED 043 033 AL 002 596

AUTHOR Angelis, Paul J.
TITLE Student Evaluation.
PUB DATE [70]
NOTE 7p.; To appear in ATEQ (Association of Teachers of English of Quebec) Journal (Special Issue No. 1, Teaching English as a Second Language), III. 2, 1970
AVAILABLE FROM Etudes Anglaises, Universite de Montreal, C.P. 6128, Montreal, Quebec

EDRS PRICE EDRS Price MF-$.25 HC-$0.45
DESCRIPTORS Instructional Materials, *Language Instruction, *Language Tests, Sampling, *Student Evaluation, Teacher Role, *Test Construction

ABSTRACT

In any field of study, teaching and learning go hand in hand with testing. In foreign language teaching, the testing of skills development has become a specialized area requiring the combined efforts of linguists, psychologists, statisticians, and educators. Teachers must constantly ask: What kind of progress are my students making in this second language course? and How effective are the instruction and the materials used in teaching what is to be learned in this course? Some language courses include series of tests; in other cases teachers must design their own tests. The most important point to keep in mind while designing tests is whether or not the results of the tests help to answer the above questions. Raw scores can be very deceiving without further information as to the difficulty of the test or how the scores of all the students relate to each other. The size of the sample must also be considered. More can be decided about the progress of a student from a number of tests taken throughout a course of instruction than from one or two tests taken at the predetermined intervals. The aim of testing is not to penalize the students. The most useful information tests should be closely connected to the teaching, and should be arranged to take the least possible amount of time. Keeping test files is useful for evaluation and review. (AMM)

ED 043033

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Student Evaluation

Paul J. Angelis
College Militaire Royal

AL 002 596

In any field of study, teaching and learning go hand in hand with testing. A student must be tested on the material which has been presented throughout a course in order to find out how much he has retained. Foreign language teaching is no exception to this rule. However, since most foreign language courses now emphasize the development of skills (mainly speaking and understanding) and not just information, the testing of these skills becomes a specialized area. Foreign language tests cannot be the same as those in history, mathematics, geography or even native languages.

Over the past few years language testing has grown into a much more important field of research and development requiring the combined efforts of linguists, psychologists, statisticians and educators. Unfortunately, this growth has served only to widen the gap between language teachers and testing specialists. Teachers see the awesome variety of test batteries along with supporting statistics as some sort of strange concoction of words, numbers and percentages by which students are supposed to be divided into neat categories in language aptitude, ability or achievement. But this is only to be expected. Why should teachers be any more up to date on the specialized information having to do with language testing than they are on similar details in educational research or linguistic analysis? Certainly, no one would discourage teachers from learning as much as they can about the background of test

development if the time is available. It is simply not a requirement.

The reason that teachers need not be concerned about a lack of detailed information on testing research becomes clear when we isolate the various types of language tests. The vast majority of formal, standardized language tests, which are the result of lengthy group research, have to do with the measurement of language aptitude (capacity for learning languages), overall proficiency (current level of language ability) or placement (how should students be divided according to this ability?). These are vital questions to administrators or those concerned with curriculum planning and student counseling but not necessarily to teachers.

The two most important questions teachers must constantly seek to answer are:

1. What kind of progress are my students making in this second language course? and

2. How effective are the instruction and the materials used in teaching what is to be learned in this course?

Both of these questions come under the general heading of evaluation. To answer them teachers can make use of tests which are designed for one course or one set of materials. Some texts come with a series of tests to be used at specified intervals in the course. One example would be the series of achievement tests which accompany An Intensive Course in English, originally prepared by Charles Fries and Robert Lado and published by The University of Michigan. In other cases teachers will need to design tests of their own.

It is with regard to these latter tests which are written, scored and interpreted by teachers for use in their own classrooms that we can search for some general guidelines. Much of the information learned in the preparation of large scale tests can apply to classroom tests as well. Perhaps the most important point to keep in mind is whether or not the results of course tests do help to answer the questions listed above. Put more specifically, does the score which student "X" received on the reading test you just gave tell you whether or not he can read that material and is making progress? The question is not as simple-minded as it looks. Raw scores (the actual scores students attain on tests) can be very deceiving without further information as to the difficulty of the test or how the scores of all the students relate to each other. Harris cites a simple example of how a raw score can give us little or no information.[1] If a group of college students were asked to recite the alphabet in proper sequence all of them would be expected to achieve a near perfect score. Even a score of 24 would seem poor. If, however, they were asked to give the same letters in reverse order at the same speed even a score of 15 might seem very good.

A second factor to be considered is the size of the sample. By this we mean not the number of students but the number of tests. More can be decided about the progress of a student if we base our judgement on a number of tests taken throughout a course of instruction than on one or two tests taken at predetermined intervals. Teachers are well aware of all of the factors which can influence a student's performance. Illness, personal problems, poor motivation and fatigue are only some of the factors which can cause a student

to do poorly on a test. If that test is the only one he has, or one of the few, we don't really have an accurate picture of what he can do. Even very short tests of no more than ten minutes duration can be useful in evaluating performance if enough of them are given throughout a semester.

But what can be done in such a short test? Probably little, unless we take time to plan what goes into the test. First of all, we must know which main points we are trying to test. This job is partly done since the test will be based on material covered in class. Most likely there will not be more than a half dozen key points covered in a two-week period. Having the language items already determined by the course does not, however, mean that the test is prepared. If the same exercises are used to test that have been used to teach there is a danger that the student will know how to use that part of the language with only those words and in only that context. We can be more certain of his command of the points to be tested if we use items which have different vocabulary and different arrangements of words.

Although language is extremely complex, it is possible to separate it into individual skills for purpose of testing. We can, for example, test only an individual's reading ability. Or, after working on the development of auditory comprehension, we can test a student on how much he understands of some selected material. The results obtained from such specialized tests are very useful for evaluation because they help to pinpoint areas where students are weak. Rarely do students of a second language master all skills

equally well. Some do better in reading than in speaking or understanding and some even do better in speaking than in understanding. Unless we know this, it is difficult to plan for individual work a student may need.

The second of the two questions we outlined earlier had to do with how well the course materials used were helping to teach what we want to teach. Let us take an example to see how tests can help us here. If a significant number of students do poorly on a section of a test having to do with the use of the past tense in English we may find that the text used has not devoted enough time to it or that the order of presentation was inadequate. The perfect language text has not been written and will most likely never be.

If we are interested in testing what we have been teaching in class then we must also realize that the scores students achieve will not necessarily fall into a neat arrangement of low and high levels with most people's scores coming in the middle. We are trying to test mastery of certain specific areas of the language we teach. Therefore we expect students to do well. The aim of testing is not to penalize them. Many psychologists and language specialists hold that strict adherence to this principle means that we would expect a 90% score as the goal in such testing. Whether we accept this or not we must admit that to give us the most useful information tests should be closely connected to the teaching we do.

Although we cannot go into any detail on the question of writing test items in this limited space, one comment can be made. Since the main job of the teacher is to teach, testing should be arranged to

take the least possible amount of time. But there remain two courses of action. We can either be quick about preparing tests which then require an undue amount of time to correct or we can try to devise tests which will give us the information we want but will be fairly easy to correct. The latter is by far the more preferable. This explains the tendency to use objective tests even on the part of professional test writers, and a number of interesting techniques have been developed which lend themselves to objective testing.[2]

Lastly, we must not neglect the importance of keeping adequate records. A teacher's file of tests need not compare with the vast store of information kept by the professional organizations which are in the business of testing. Nevertheless, it is extremely helpful to know how tests have been used in the past. In order to fully answer our second question on the merits of materials, we may have to keep records for a number of terms or even years. From the point of view of the student also, records help to review long-range developments in the progress he has made and maybe even prospects for the future.

In summary then, we cannot stress too much the close connection between teaching and testing. Notwithstanding the continuing research in language testing, there is much that the individual teacher can do in language testing, there is much that the individual teacher can do in planning, administering and interpreting the tests he uses in his classes. This is his job; the research belongs to the specialists. But we cannot deny that a two-way line of communication exists between

teachers and researchers. What teachers find in using tests will ultimately determine the courses of action followed by testing specialists. Even now the initial test items used in the preparation of large-scale standardized tests are submitted by language teachers. May this always be the case.

Footnotes

1. David P. Harris, Testing English as a Second Language, New York: McGraw Hill, 1969. p. 121.

2. Many helpful suggestions can be found in the following works:

   Robert Lado, Language Testing: The Construction and Use of Foreign Language Tests, New York: McGraw Hill, 1961.

   Rebecca M. Valette, Modern Language Testing: A Handbook, New York: Harcourt, Brace & World, 1967.